7.5/7.5

Steve Depp
413 – 55

Assignment 6
22 July 2019

Report needs a forecast plot

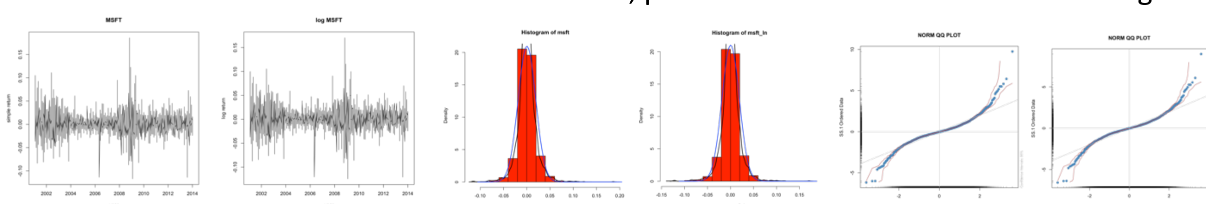
1.1 Use EDA to justify a transformation of the simple returns to log returns.

The following presents sample size, mean, standard deviation, median, 20% trimmed mean, min, max, range, skewness, excess kurtosis for MSFT simple and log returns over 3269 trading days (3268 returns). EDA first confirms there are no gaps in the 605 monthly observations and converts dates from YYYYMMDD format to monthly percent of year format.

	n	mean	sd	median	trimmed	min	max	range	skew	kurtosis
msft	3268	0.0004403911	0.01903683	0	0.0002242748	-0.1171310	0.1860470	0.3031780	0.4125465	7.622458
msft_ln	3268	0.0002599197	0.01898513	0	0.0001818206	-0.1245784	0.1706259	0.2952044	0.1526723	6.966618

The **above** univariates describe distributions with near zero mean departing from normal via leptokurtosis, positive skew. Although the median is exactly zero, the range of distribution extends further to the right than the left. From positioning of median < trimmed < mean, the middles of the distributions are also skewed right.

All simple return univariates are less than the same log return univariates. Log transformation attenuates non-zero mean, positive skew and kurtosis to some degree.



In the above pairs of plots, log returns (on the right) are less than simple returns. This is most evident in the axes of the histogram and QQ plots which extend to 0.20 and 10 SDs for simple returns but only 0.15 and 5 SDs for log returns. Otherwise, the plots suggest no benefit to log transformation. Both distributions are highly leptokurtotic and somewhat right skewed.

Is the expected log return zero? Why?

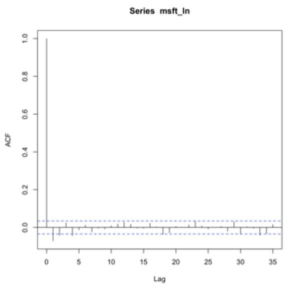
```
[1] "msft H1: mean=zero"

        One Sample t-test

data:  msft
t = 1.3225, df = 3267, p-value = 0.1861
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 -0.0002125329  0.0010933151
sample estimates:
   mean of x
0.0004403911
```

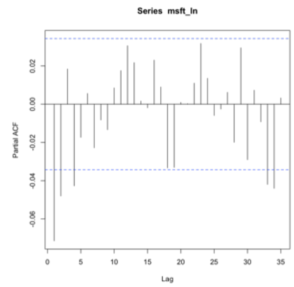
```
[1] "msft_ln H1: mean=zero"

        One Sample t-test

data:  msft_ln
t = 0.78265, df = 3267, p-value = 0.4339
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 -0.0003912310  0.0009110703
sample estimates:
   mean of x
0.0002599197
```

The p-value = 0.1861 and 0.4339 confirms respectively that one cannot reject the null hypothesis of zero mean for both distributions. For both distributions, test statistics allow one to reject the hypothesis of zero excess kurtosis at the 99.9% confidence level. Similarly, for skew, the test statistic allows one to reject the hypothesis of zero skew in both distributions. The expected log return is zero because the sample mean of log returns is statistically no different from zero with 99.9% confidence. Beyond the sampling distribution, it is reasonable for the expected log return to be close to zero for some

⇒ multiplicative symmetry as now add the stocks and not for others. Stocks that pay no dividends would tend to have positive mean returns to compensate for no dividend yield. (Not sure if this was the crux of your question.)

Are there any serial correlations in the log returns? Why ?






```
Box-Ljung test

data:  msft_ln
X-squared = 34.476, df = 10, p-value = 0.0001534
```

In the above ACF and PACF plots there are significant autocorrelations at lag 1, 2, and 4. The PACF autocorrelations decline exponentially while the ACF are relatively level and cut off after lag 4. It is possible that any or all of these might be due to chance alone, but where these indicative of truly non-zero serial correlation, they might indicate that an ARMA (0, 4) model with a zeroed 3rd coefficient would be worth testing. The Box-Ljung portmanteau test suggests non-zero autocorrelation inside the 10 lages at the 99.9% confidence level.

1.2 The following presents a MA(4) model to be fitted withoout mean:

$$\ln(r_t) = (1+\theta_1 B+ \theta_2 B^2+ \theta_3 B^3+ \theta_4 B^4)\varepsilon_t$$ $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$

$$\ln(r_t) = a_t+\theta_1 a_{t-1}+ \theta_2 a_{t-2}+ \theta_3 a_{t-3}+ \theta_4 a_{t-4}$$

The fitted MA(4) mean model is

$$\ln(r_t) = a_t + (-0.0754)\, a_{t-1}+ (-0.0424)\, a_{t-2}+ (0.0183)\, a_{t-3}+ (-0.0448)\, a_{t-4}$$

with standard errors 0.0176, 0.0176, 0.0174, and 0.0176 for the parameters.

The starter code MA(2) model fitted without mean would be:

$$\ln(r_t) = a_t+\theta_1 a_{t-1}+ \theta_2 a_{t-2}$$

when fit is: $\ln(r_t) = a_t + (-0.074)\, a_{t-1}+ (-0.0444)\, a_{t-2}$

```
Box-Ljung test

data:  m1r2
X-squared = 603.99, df = 10, p-value < 2.2e-16
```

The above results a Box-Ljung test of the squared residuals of the fitted MA(2) model with Q(10) = 603.99 has a p-value < 2.2e-16 indicating we can reject the null hypothesis of zero serial correlation indicating strong ~~G~~ARCH effects are present.

1.3 The joint estimate of a an MA(2)-GARCH(1,1) model would have this model to be fitted:

$$\ln(r_t) = a_t+\theta_1 a_{t-1}+ \theta_1 a_{t-2} + \beta_0$$
$$\sigma_t^2= \alpha_0 +\beta_1\sigma_{t-1}^2 + \alpha_1 a^2_{t-1}$$

?

when fit yields:

$$\ln(r_t) = a_t+(-0.043928)a_{t-1}+ (-0.0014333)a_{t-2} + 0.00037801$$
$$\sigma_t^2= 0.000004396 + 0.93494\sigma_{t-1}^2 + 0.051724a^2_{t-1}$$

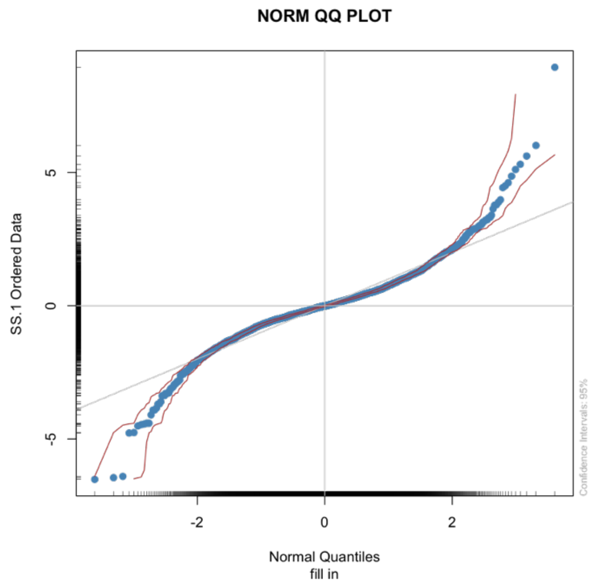
```
        mu          ma1          ma2        omega       alpha1        beta1
 3.7801e-04  -4.3928e-02  -1.4333e-03   4.3060e-06   5.1724e-02   9.3494e-01

Std. Errors:
 based on Hessian

Error Analysis:
        Estimate  Std. Error  t value Pr(>|t|)
mu      3.780e-04   2.471e-04    1.530   0.1261
ma1    -4.393e-02   1.896e-02   -2.316   0.0205 *
ma2    -1.433e-03   1.949e-02   -0.074   0.9414
omega   4.306e-06   1.074e-06    4.010 6.07e-05 ***
alpha1  5.172e-02   9.428e-03    5.486 4.11e-08 ***
beta1   9.349e-01   1.174e-02   79.625  < 2e-16 ***
```

The parameters' standard errors shown above indicate the second parameter MA1 is significant only at the 95% confidence level (p-value = 0.0205), that the 3rd parameter MA2 may not be significant at all (p-value = 0.94) and mu may not be different from zero: pvalue = 0.1261 indicates we cannot reject the null hypothesis that mu or intercept of the MA model is different from zero, which is consistent with our original test of mean returns.

```
Standardised Residuals Tests:
                                Statistic p-Value
 Jarque-Bera Test   R    Chi^2  13248.56  0
 Shapiro-Wilk Test  R    W      0.9445296 0
 Ljung-Box Test     R    Q(10)  8.162707  0.6129475
 Ljung-Box Test     R    Q(15)  17.05979  0.3153029
 Ljung-Box Test     R    Q(20)  21.46915  0.3699963
 Ljung-Box Test     R^2  Q(10)  0.5607348 0.9999886
 Ljung-Box Test     R^2  Q(15)  2.350846  0.9999144
 Ljung-Box Test     R^2  Q(20)  2.689185  0.9999984
 LM Arch Test       R    TR^2   0.7208459 0.9999978
```

Above, the model's Ljung-Box tests' Q(m) for m = 10 through 20 indicates no serial correlation among the model residuals or their squares. The LM Arch test with a test statistic 0.72 and pvalue 0.99 confirms the Ljung-Box test of residual squares that no

GARCH effects are left in the residuals. So it appears to be a good fit.

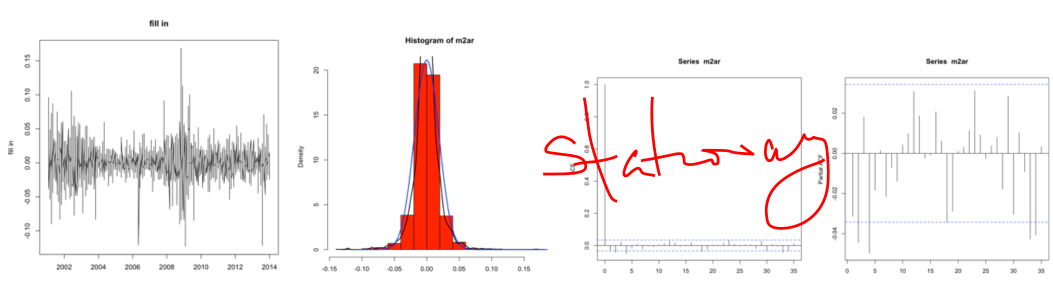


The QQ plot of the standardized residuals shown above exhibit fat tails and likely moderate positive skew given the extreme outlier near 10 SDs and though outliers beyond 5 SDs appear in equal numbers. The left side of the distribution departs from normal earlier and more severely but looking at the residuals numerically will tell.

```
[1] "m2br H1: mean=zero"

        One Sample t-test

data:  m1ar
t = 0.91687, df = 3267, p-value = 0.3593
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 -0.0003449701  0.0009509932
sample estimates:
   mean of x
0.0003030116
```

	n	mean	sd	median	trimmed	min	max	range	skew	kurtosis
m2ar	3268	-0.0001588802	0.01885853	-0.0003778328	-0.0002038666	-0.1230971	0.1685167	0.2916137	0.1048228	6.834280

The univariate statistics of residuals for the model show similar degrees of skew and kurtosis as the original log return data set and the p-value = 0.3593 from a one sample t-test of the mean would not allow us to reject the null hypothesis that the mean = 0. Similar tests for kurtosis and skew confirm they are significant at the 99.9% confidence level.



Stationary

The residuals time series plot, histogram and ACF plots exhibit stationarity, normality, and a lack of auto correlation respectively. The PACF plot however shows significant partial autocorrelation at the 2nd and 4th lag which could be due to randomness of a 95% confidence level.

Independence?

A test for unit root for the residuals via Dickey Fuller indicated we can reject the hypothesis of unit roots / non-stationarity (ADF). A test of the same via KPSS suggests we cannot reject the null hypothesis of stationarity of residuals.

```
 Augmented Dickey-Fuller Test

Test Results:
  PARAMETER:
    Lag Order: 5
  STATISTIC:
    Dickey-Fuller: -16.3787
  P VALUE:
    0.01
```

```
#######################
# KPSS Unit Root Test #
#######################

Test is of type: mu with 9 lags.

Value of test-statistic is: 1.2998

Critical value for a significance level of:
                10pct  5pct 2.5pct  1pct
critical values 0.347 0.463  0.574 0.739
```

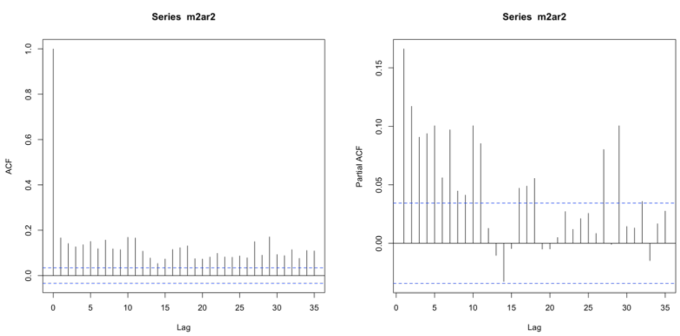
Odd results

For squared residuals however, the KPSS test **above on the right** can be rejected at the 99% level with a test statistic of 1.2998 and repeated tests via the ndiff function finds that there is room for a 1st difference in the squared residuals.

```
        Box-Ljung test

data:  m2ar
X-squared = 21.828, df = 10, p-value = 0.016
```

```
        Box-Ljung test

data:  m2ar2
X-squared = 652.91, df = 10, p-value < 2.2e-16
```

A check of Box-Ljung residuals and residuals squared **above** yielded pvalues that were very close to zero.



ACF and PACF of squared residuals exhibit strong serial correlation throughout and strong partial autocorrelation to lag 11. These and the individual Box-Ljung tests would contradict the model output. So, I might have to look at these a 4th time (already checked a few times).

Although the model appears to be adequate in describing the linear dependence in the return and volatility series, the fitted model shows shows $\hat{\alpha}_1 + \hat{\beta}_1 = 0.9349 + 0.0517 = 0.9866$, which is close to 1 which leads to imposing the constraint $\alpha_1 + \beta_1 = 1$ in a GARCH(1,1) model, resulting in an integrated GARCH (or IGARCH) model to be explored in problem 1.6.

All the same conclusions reached above hold for the joint model that leaves out the MA2 parameter, which is the starter code's other joint model, MA(1)-GARCH(1,1), that will be used in the next problem.

1.4. A joint estimate of ARMA-GARCH model with Student-t innovations would fit this model MA(1)-GARCH(1,1):

$$\ln(r_t) = a_t+\theta_1 a_{t-1} + \beta_0$$
$$\sigma_t^2 = \alpha_0 + \beta_1\sigma_{t-1}^2 + \alpha_1 a^2_{t-1}$$

When fit we obtain:

$$\ln(r_t) = a_t + (-0.04457)a_{t-1} + 0.000288$$
$$\sigma_t^2 = 0.000001571 + 0.9412\sigma_{t-1}^2 + 0.05627a^2_{t-1}$$

with estimated degrees of freedom = 4.797

```
Coefficient(s):
        mu         ma1       omega      alpha1       beta1       shape
 2.2882e-04  -4.4567e-02  1.5707e-06  5.6271e-02  9.4123e-01  4.7971e+00

Std. Errors:
 based on Hessian

Error Analysis:
         Estimate  Std. Error  t value Pr(>|t|)
mu      2.288e-04   2.086e-04    1.097  0.27266
ma1    -4.457e-02   1.691e-02   -2.636  0.00838 **
omega   1.571e-06   6.058e-07    2.593  0.00952 **
alpha1  5.627e-02   9.809e-03    5.737 9.65e-09 ***
beta1   9.412e-01   9.496e-03   99.115  < 2e-16 ***
shape   4.797e+00   3.891e-01   12.330  < 2e-16 ***
```

All parameters above are significant at at least the 99% confidence level except mu

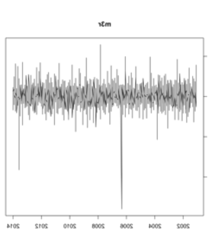
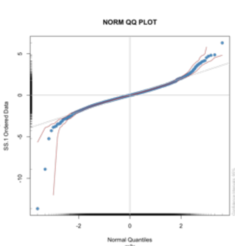
```
Standardised Residuals Tests:
                                Statistic p-Value
 Jarque-Bera Test   R    Chi^2  28541.15  0
 Shapiro-Wilk Test  R    W      0.9338294 0
 Ljung-Box Test     R    Q(10)  8.835501  0.5477832
 Ljung-Box Test     R    Q(15)  17.79975  0.2733399
 Ljung-Box Test     R    Q(20)  22.10929  0.3346098
 Ljung-Box Test     R^2  Q(10)  0.6060827 0.9999834
 Ljung-Box Test     R^2  Q(15)  2.153636  0.9999517
 Ljung-Box Test     R^2  Q(20)  2.62604   0.9999987
 LM Arch Test       R    TR^2   0.994827  0.9999862
```

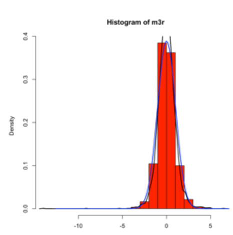
The **above** tests of standardized Student-t residuals and their squares indicate no serial correlation for Ljung-Box tests Q(m) with m=10, 15, 20 and for Lagrange Multiplier tests for conditional heteroscedasticity.

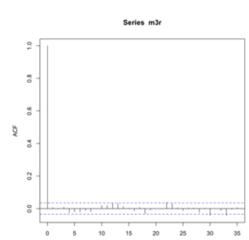
	n	mean	sd	median	trimmed	min	max	range	skew	kurtosis
m3r	3268	0.001815629	1.024579	-0.01360255	-0.004784957	-13.98175	6.452228	20.43398	-0.7264739	14.39392
m3r2	3268	1.049444535	4.250489	0.31599609	0.533962483	0.00000	195.489362	195.48936	32.3403670	1384.56169

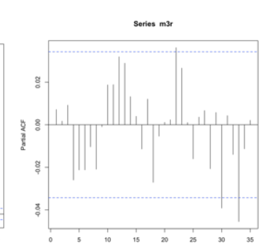
```
[1] "m3br H1: mean=zero"

        One Sample t-test

data:  m3r
t = 0.1013, df = 3267, p-value = 0.9193
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 -0.03332532  0.03695658
sample estimates:
  mean of x
0.001815629
```

Univariates for the residuals and their squares are shown above together with a one sample t-test of residuals mean indicating we cannot reject the null hypothesis that mean = 0. The residuals are negatively skewed and leptokurtotic at a confidence level of 99.9%. One can see that although mean residuals of the model are zero, the range of residuals mostly stretches toward negatives with a min = -14 and max 6.5. From the proximity of trimmed versus mean, it appears the negative skew is mostly in the wings of the distribution.

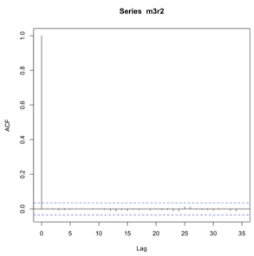



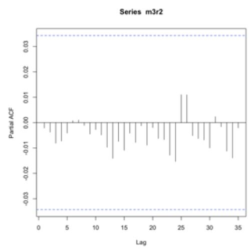






Time series of residuals, QQ plot and histogram **above** bear out the conclusions from univariate measures that the distribution is zero meaned, very leptokurtotic and negatively skewed with 2 outliers near 8 and 15 standard deviations from the mean. ACF and PACF bear out stationarity with no significant auto or partial auto correlations out to lag 20.





The tests of residuals and squared residuals for unit roots via Dickey Fuller, KPSS yielded results confirming stationarity. Individual Box-Ljung tests on residuals and squared residuals are consistent with the model output. Residual squared ACF and PACF plots also are consistent with the model capturing all the serial correlation in the conditional variance. That is there are no remaining Garch effects.

The model appears to be adequate in describing the linear dependence in the return and volatility series, but again the fitted model shows shows $\hat{\alpha}_1 + \hat{\beta}_1 = 0.9412 + 0.0563 = 0.9975$, which is close to 1 which leads to imposing the constraint $\alpha_1 + \beta_1 = 1$ in a GARCH(1,1) model, resulting in an integrated GARCH (or IGARCH) model to be explored in problem 1.6.

meanForecast	meanError	standardDeviation
9.662615e-05	0.01222158	0.01222158
2.288235e-04	0.01228253	0.01227045
2.288235e-04	0.01233113	0.01231900
2.288235e-04	0.01237942	0.01236724
2.288235e-04	0.01242740	0.01241517

1.5

The **above** table presents the 1-step and 5-step ahead forecasts of the ARMA-GARCH model with Student-t innovations.

```
Coefficient(s):
       Estimate  Std. Error  t value   Pr(>|t|)
beta 0.97306667  0.00253629  383.658 < 2.22e-16 ***
```

1.6

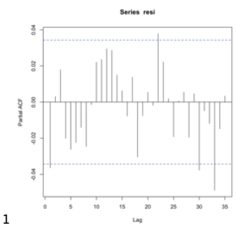
IGARCH model to be fitted are **below** and the fitted Beta and p-value are **above**:

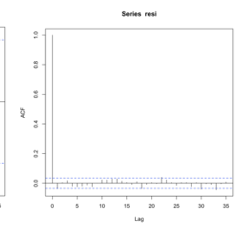
$$\ln(r_t) = a_t$$
$$\sigma_t^2 = \alpha_0 + \beta_1\sigma^2_{t-1} + (1 - \beta_1)a^2_{t-1}$$

Fit model equation is thusly: $\sigma_t^2 = 0.973\sigma^2_{t-1} + (0.027)a^2_{t-1}$

```
        Box-Ljung test

data:  resi
X-squared = 14.48, df = 10, p-value = 0.1522
```

```
        Box-Ljung test

data:  resi2
X-squared = 0.47788, df = 10, p-value = 1
```

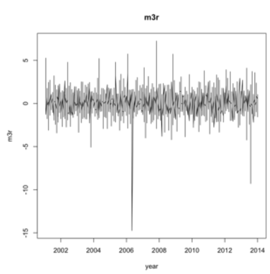


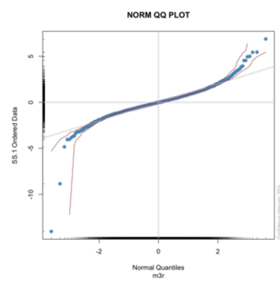


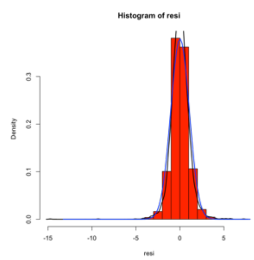
1.7 and 1.8

Neither the **above** Box-Ljung test of residuals and squared residuals of the IGARCH model nor the residuals ACF and PACF indicate any serial correlation as the unit root of the series has removed.

	n	mean	sd	median	trimmed	min	max	range	skew	kurtosis
X1	3268	0.01589438	1.05208	0	0.008356278	-14.74036	7.250053	21.99041	-0.7300119	15.97266







The above residual time series, QQ and histogram present the same stationary series that is highly skewed left by a single outlier, otherwise normal in appearance.
The model has performed the equivalent of 1st differencing to remove the unit root and any serial correlation with it.

1.9

Appears the model is adequate and feeding the model values results in volatilities that are steady:

0.0138965075044220
0.0138969644983408
0.0138969440918393
0.0138969242280626

1.10

Preference for the MA(1)-GARCH(1,1) employing Student-t distribution model because it's AIC is -5.51 vs -5.37 for the gaussian distribution, and at this point because I am unable to fully appreciate the IGARCH model. The original MA(2)-GARCH(1,1) model needed to be reduced to one MA parameter and so is not considered.

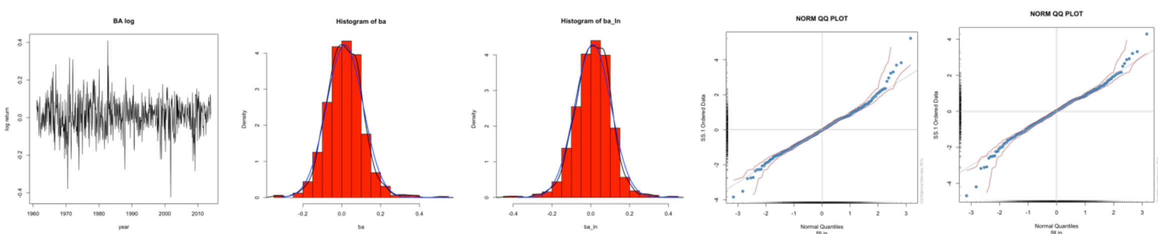
```
        One Sample t-test

data:  ba
t = 4.0722, df = 635, p-value = 5.245e-05
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 0.007867071 0.022520642
sample estimates:
 mean of x
0.01519386
```

```
        One Sample t-test

data:  ba_ln
t = 2.9259, df = 635, p-value = 0.003557
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 0.003547215 0.018025521
sample estimates:
 mean of x
0.01078637
```

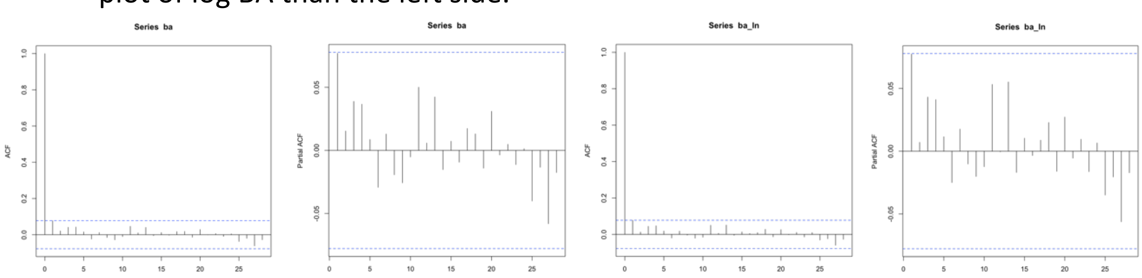
2.1 The expected log return of log BA is not zero. The one sample t-test of log returns shown **above on the right** has a p-value of 0.003557 indicating we can reject the null hypothesis that the mean is zero at the 99% confidence level.

	n	mean	sd	median	trimmed	min	max	range	skew	kurtosis
ba	636	0.01519386	0.09409475	0.01308600	0.01388385	-0.3457030	0.5062500	0.851953	0.3062674	1.897722
ba_ln	636	0.01078637	0.09296932	0.01300112	0.01212786	-0.4241939	0.4096231	0.833817	-0.2003728	1.683996

Above are the univariates of BA and log BA. Comparing the two series, one can see the distribution shifts left ward when log transformed, and thus positive skew becomes negative skew. The mean shifts leftward as well, but we are only marginally less confident that we can reject the null hypothesis of mean zero with log returns. As will be seen with univariate plots, the skew shift toward the left (negative) with log returns also makes us slightly less confident of rejecting the null hypothesis of zero skew with log returns: pvalues for a skew test is 0.0016 with BA and 0.039 with log BA. So there is some benefit seen in univariate measures for log transformation.



One can clearly see a less skewed distribution in the right hand log BA versus left hand BA histogram plots **above.** It is difficult to see any change in the time series; so, only one is shown to demonstrate the placement of outliers through time, which one can see in the QQ plot. The QQ plot above, as was the case with MSFT, only shows a shift in skew due to the shift in the y-axis more evenly spread around the data on the right side plot of log BA than the left side.



There is no serial correlation evident in the ACF / PACF plots of either BA or log BA series **above**. A Box-Ljung test with Q(12) has a pvalue of 0.66 indicating no serial correlation out to lag 12. Setting the model equation

$$a_t = \log(r_t) - mu$$

and squaring the residuals of this model, then testing for serial correlation Q(12) of these squared residuals reveals that we can reject the null hypothesis of no serial correlation with 99% confidence (p-value = 0.001305). Thus, there are GARCH effects in the log returns.

2.2
The model to be fitted has the mean equation suggested above plus a GARCH(1,1) variance equation:

$$\ln(r_t) = a_t + \beta_0$$
$$\sigma_t^2 = \alpha_0 + \beta_1\sigma_{t-1}^2 + \alpha_1 a^2_{t-1}$$

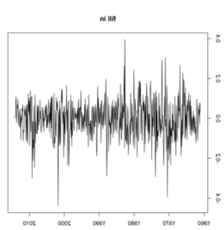
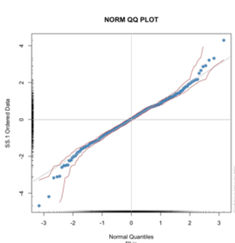
When fit, the model is:

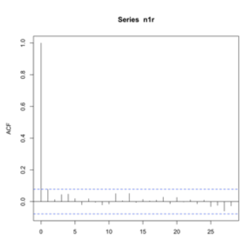
$$\ln(r_t) = a_t + 0.0148031$$
$$\sigma_t^2 = 0.0002526 + 0.8441071\sigma_{t-1}^2 + 0.1362668a^2_{t-1}$$

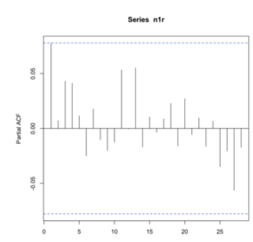
	n	mean	sd	median	trimmed	min	max	range	skew	kurtosis
n1r	636	-0.004016715	0.09296932	-0.001801968	-0.002675229	-4.389970e-01	0.3948200	0.8338170	-0.2003728	1.683996
n1r2	636	0.008645838	0.01670437	0.003609202	0.005223757	1.353604e-10	0.1927184	0.1927184	5.6291643	44.470010

```
        One Sample t-test

data:  n1r
t = -1.0896, df = 635, p-value = 0.2763
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 -0.011255868  0.003222437
sample estimates:
   mean of x
-0.004016715
```

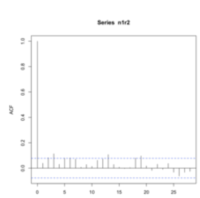
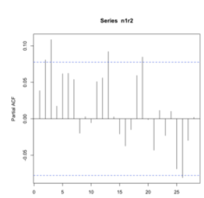
The residuals of the model are negatively skewed, leptokurtotic with mean whose p-value = 0.2763 would not enable rejection of the null hypothesis that the mean = 0. The negative skew of residuals is statistically significant at the 95% level (p-value = 0.039) enabling rejection of the null hypothesis. Positive excess kurtosis is also significant at the 99.90% confidence level.





Time series of residuals **above** appear stationary with near zero mean and QQ is skewed and leptokurtotic with outliers. ACF and PACF exhibit no significant serial correlation. There are bunches of higher and lower variance evident in the time series, but it would be difficult to estimate a pattern without examining the squared residuals ACF and PACF.

```
        Box-Ljung test

data:  n1r2
X-squared = 32.335, df = 12, p-value = 0.001228
```

There are 3 – 4 lags (2,3,13, 19) inside of 20 that exhibit significant partial serial correlation of the squared residuals in the right hand plot above. Possibly 5 lags (2,35,6,7, 13) have borderline significant autocorrelation in the ACF on the left side **above**. While the Box-Ljung test of residuals allowed rejection of portmanteau set of significant autocorrelations inside lag 12, the same test for squared residuals yielded a test statistic that enables us to reject the hypothesis of zero autocorrelation inside 12 lags with 95% confidence (p-value = 0.001228) as shown **above**. This means there are GARCH effects in this series.

```
Coefficient(s):
        mu       omega      alpha1       beta1        skew       shape
1.2604e-02  2.9147e-04  1.0801e-01  8.6368e-01  8.8818e-01  1.0000e+01

Std. Errors:
 based on Hessian

Error Analysis:
        Estimate  Std. Error  t value Pr(>|t|)
mu     1.260e-02   3.350e-03    3.763 0.000168 ***
omega  2.915e-04   1.451e-04    2.008 0.044632 *
alpha1 1.080e-01   3.049e-02    3.542 0.000396 ***
beta1  8.637e-01   3.359e-02   25.714  < 2e-16 ***
skew   8.882e-01   5.998e-02   14.807  < 2e-16 ***
shape  1.000e+01   2.838e+00    3.524 0.000425 ***
```

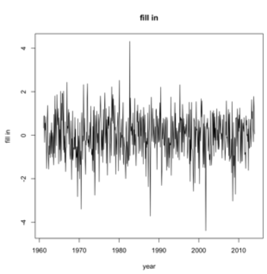
2.3 The model output **above** is from the fitted model which has the mean and variance equations **below**:

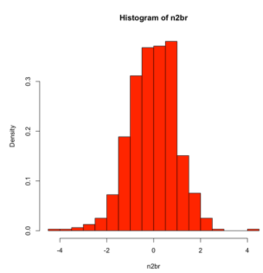
$$\ln(r_t) = a_t + \beta_0$$
$$\sigma_t^2 = \alpha_0 + \beta_1\sigma_{t-1}^2 + \alpha_1 a^2_{t-1}$$

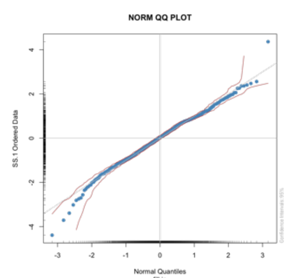
When fit, the model is as shown **here**:

$$\ln(r_t) = a_t + 0.012604$$
$$\sigma_t^2 = 0.00029147 + 0.86368\sigma_{t-1}^2 + 0.10801 a^2_{t-1}$$

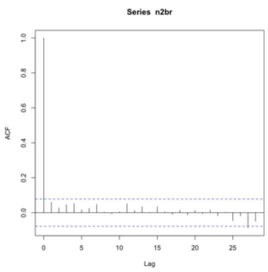
	n	mean	sd	median	trimmed	min	max	range	skew	kurtosis
n1r	636	-0.05023973	0.9971571	-0.018432113	-0.034271839	-4.290774	4.234862	8.525637	-0.2071900	0.8802270
n2ar	636	-0.03891011	0.9925008	-0.007018312	-0.022179200	-4.369595	4.266054	8.635649	-0.2253958	0.9885517
n2br	636	-0.02523503	0.9940480	0.003584221	-0.008153593	-4.387859	4.312517	8.700376	-0.2300169	1.0269055

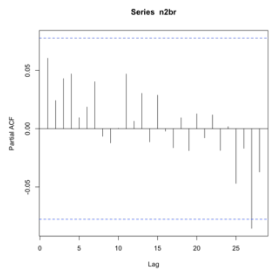


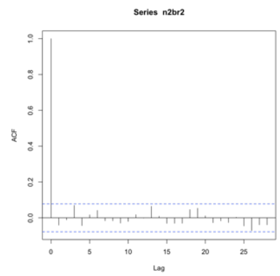


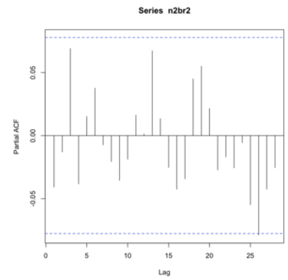


Standardized residual univariates without Students-t, with Students-t, with skewed Students-t are shown above and are not terribly different. If anything, the latter 2 are more kurtotic and negatively skewed. This additional skew is only slightly more evident in the above time series, histogram and QQ plot, all 3 of which still appear to be adequate. The time series appears stationary and a single t-test shows that a null hypothesis that its sample mean = 0 cannot be rejected.









The ACF and PACF for residuals and their squares also evidence a model which has captured all residual serial correlation and the GARCH effects since all auto and partial auto correlations are within their confidence bands.

```
        Box-Ljung test

data:  n2br
X-squared = 9.8122, df = 12, p-value = 0.6324


        Box-Ljung test

data:  n2br2
X-squared = 8.2133, df = 12, p-value = 0.7682
```

Box-Ljung tests for both residuals and their squares also evidence no autocorrelations inside of lag 12.

Given that the model is adequate and the parameter for skew has a p-value = 0, it would appear that the stocks' log returns are skewed.

```
Coefficient(s):
         Estimate  Std. Error  t value   Pr(>|t|)    
mu    0.011491820 0.006786824  1.69325   0.090407 .  
gamma 0.956124331 0.797836314  1.19840   0.230763    
omega 0.000250370 0.000128816  1.94363   0.051940 .  
alpha 0.129916458 0.032094830  4.04789 5.1681e-05 ***
beta  0.850389657 0.032169897 26.43433 < 2.22e-16 ***
```

2.4 The GARCH M model to be fit would have the equation

$$\ln(r_t) = a_t + c\,\sigma_t^2 + \beta_0$$
$$\sigma_t^2 = \alpha_0 + \beta_1\sigma_{t-1}^2 + \alpha_1 a^2_{t-1}$$

When fit, as shown above, the model is as shown **here**:

$$\ln(r_t) = a_t + 0.9561\,\sigma_t^2 + 0.01149$$
$$\sigma_t^2 = 0.00025037 + 0.85038\sigma^2_{t-1} + 0.12992 a^2_{t-1}$$

The p-value for the gamma coefficient is 0.23 indicating that we cannot reject the null hypothesis that this coefficient is zero. (It has a large standard error.). Thus it is unlikely the risk premium component of mean model is statistically significant.

2.5 The TGARCH model to be fit would have the equation

$\ln(rt) = at + \beta_0$
$\sigma_t^2 = \alpha_0 + \beta_1 \sigma_{t-1}^2 + (\alpha_1 + \gamma_i N_{t-1})a^2_{t-1}$

$$N_{t-i} = \begin{cases} 1 & \text{if } a_{t-i} < 0, \\ 0 & \text{if } a_{t-i} \geq 0, \end{cases}$$

When fit, as shown above, the model is as shown **here**:

$\ln(r_t) = a_t + 0.016049$
$\sigma_t^2 = 0.000241725 + 0.8496\, \sigma^2_{t-1} + (0.13307 + 0.13307 * N_{t-1})\, a^2_{t-1}$

The parameter omega which scales the threshold is significant at the 95% confidence level based upon the p-value = 0.0418.

3.0 Boeing stock rises an average of 1% per month with 9% stock volatility. Over the course of history, the range of monthly returns have been +/- 40% and generally there is no positive return rewarded to holders from bearing additional volatility risk. There is however some threshold for leverage that is significant so that volatility rises or falls quickly as returns pass that threshold. Presently we have tested for zero threshold but knowing this threshold will be useful for knowing when to enter and exit the stock to obtain lower or avoid higher volatile periods respectively.

forecast plot?